

METHOD AND APPARATUS FOR LOCATING CONTENT IN A PROGRAM

TECHNICAL BACKGROUND

5 The present invention relates to a method and apparatus for locating program contents, particularly to a method and apparatus for locating according to the contents of the multimedia programs.

10 In addition to a video stream and an audio stream, a multimedia program generally contains an image stream and/or a text stream, these streams are synchronized with each other according to particular rules and predetermined time sequence for users to enjoy. Among the numerous multimedia program editing rules, the synchronized multimedia integration language (SMIL) is a popular editing language. The SMIL can not only integrate the respective content streams of a multimedia program in time
15 sequence, but also be used to manage the layout of the multimedia program being presented.

20 While watching a multimedia program, a user sometimes needs to find a particular segment of the program. For example, one needs to find the part on Iraq within a multimedia program of the lecture given by President Bush at Tsinghua University. The user can identify the audio content by fast winding/rewinding the recording media to locate the parts in the program. For another example, a user expects to directly browse a segment on the Sydney Opera Theater in a multimedia recording program depicting Australian scenery. To meet this need, the multimedia playing
25 apparatus should be able to automatically matching analysis the contents of the video streams, so that when the Sydney Opera Theater appears, the related segment is presented to the user.

 In the process of content locating as described above, if a user performs location manually, he/she has to perform repeatedly the search

before finding the desired position of the segment, which would be time-consuming and bothersome. If the user performs the locating by auto-scanning the multimedia playing apparatus, the search will be very difficult due to the complexity and magnitude of video and audio streams, more sophisticated hardware will be needed, which would increase the cost for the user.

Additionally, there are various authoring tools, such as the PresenterOne developed by the US Accordent Corporation and the Canadian Presentation Maker developed by the SofTV.net Corporation in the market for the convenience of editing multimedia programs, especially for editing multimedia demonstration programs. These tools allow a user to enlist the titles of text slides of multimedia demonstration, and the user can use these titles as indexes to locate corresponding segments. Although it simplifies the searching process to a certain degree, the above editing tools must be used in making the multimedia demonstration program. Furthermore, the editing tools provide only a very limited number of titles for users to choose from, which restricts the arbitrary of the user's choices, and renders the user-based choice impossible.

Therefore, a new program content locating method and apparatus is needed, which enables users to locate program contents in multimedia programs conveniently so that their individual requirements could be satisfied by obtaining any segments as they want.

SUMMARY OF THE PRESENT INVENTION

Therefore, one of the objects of the present invention is providing a new program content locating method and apparatus to overcome the defects of the prior art, which enables users to locate program contents in multimedia programs conveniently to obtain the particular segments as they want.

The present invention provides a method for locating content in a multimedia program, which comprising a stream with word symbol information, comprising: firstly receiving a request comprising a specific word symbol from a user; then determining a position where the specific word symbol appears in the stream with word symbol information; and finally
5 determining other presentable information synchronous with the word symbol information at the position. The other presentable information may be video information or audio information.

The word symbol information may exist in a text format or image
10 format. When it exists in an image format, the locating method further comprises the step of obtaining the text information corresponding to the word symbol information.

The stream provided with word symbol information may have a layered structure. If so, the locating method further comprises the step of
15 determining a layer containing the position where specific word symbol appear and having a particular starting position and a particular end position, so that the other finally determined presentable information has the corresponding start position and end position.

The present invention further provides an apparatus for locating
20 contents in a multimedia program, which has a stream provided with word symbol information. The word symbol information may exist in a text format or an image format. The apparatus includes a request receiving means, a word symbol locating means and a synch-locating means.

The request receiving means is used to receive a request comprising
25 a specific word symbol from a user; the word symbol locating means is used to determine the position where the specific word symbol appear in the stream provided with word symbol information; and the synch-locating means is used to determine the other presentable information that synchronizes with the word symbol information appearing at the position.

The other presentable information may be video information or audio information.

The present invention locates the position of a user required segment in a program by analyzing the stream provided with word symbol information which is included in multimedia programs, then finds corresponding video or audio segment according to the synchronization rules. Since the streams provided with word symbol information, such as text streams or image streams, contain much less a quantity of data relative to video or audio, and the analysis of text is also much simpler than that of picture or audio; therefore, the present invention has greatly reduced the complexity of searching program contents, lowered the hardware requirement, made user's operation convenient, and satisfied different needs of individual users.

The other objects and advantageous of the present invention will be evident, and the present invention also be better understood through the description and claims made with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is explained in detail by way of examples and with reference to the accompanying drawings, in which:

Figure 1 is a system block diagram of an apparatus for locating contents in a multimedia program according to an embodiment of the present invention;

Figure 2 is a flow chart of the process for locating contents in a multimedia program according to an embodiment of the invention;

Figure 3 is a flow chart of the process for locating contents in a multimedia program and extracting particular segments according to another embodiment of the present invention;

In all the figures, the same reference numbers indicate similar or identical features and functions.

DETAILED DESCRIPTION OF THE EMBODIMENTS

Figure 1 shows a system block diagram of an apparatus for locating contents in multimedia programs according to an embodiment of the present invention. The apparatus 100 may be part of a multimedia program making apparatus (not shown in this figure) or a multimedia playing apparatus (not shown in this figure). The apparatus 100 includes a request receiving module 120, a text locating module 130 and a synch-locating module 140. The apparatus 100 further includes a content receiving module 110, a presentation module 150 and an extraction module 160. Said module included in the apparatus 100 can be realized by those skilled in the art by the various existing module as long as their combination can perform the functions of the present invention.

The content receiving module 110 is used to receive a multimedia program, which contains a stream provided with word symbol information, such as a text stream or an image stream having the word symbol information (as slides of the auxiliary demonstration tools in existing multimedia demonstration programs, e.g., one page of a PowerPoint file, sometimes transmitted in an image format). The multimedia program may come from a local storage module (not shown in the figure), such as a DVD, or from a web server (not shown in the figure).

Request receiving module 120 is used to receive a request, which contains specific word symbol, such as, "Sydney Opera Theater". A user hopes to find with this request in the segment on Sydney Opera Theater in the multimedia program being edited/appreciated. The multimedia program includes a stream provided with word symbol information.

Text locating module 130 is used to determine the position of specific word symbol in the multimedia program. Module 130 searches the specific word symbol, such as "Sydney Opera Theater", in the stream provided with word symbol information, and, after the specific word symbol are found,

obtains the information on their positions in the programs. If the stream provided with word symbol information is an image stream, the module 130 is further used to obtain the text information corresponding to the word symbol information in the image stream.

5 Synch-locating module 140 is used to determine the other presentable information that synchronizes with the word symbol information appearing at the site. Because of the synchronization in time of different content streams in a multimedia program, it is possible to determine the corresponding positions of a site in the other content streams, such as
10 video streams or audio streams, according to the position information about the site in one content stream, such as a text stream.

Presentation module 150 is used to present to the user the program contents in the particular position within a multimedia program.

15 Extracting module 160 is used to extract a particular segment from a multimedia program. In this embodiment, the particular segment may contain the particular text information.

For the details of the flow chart showing the operation running module 100, see Figures 2 and 3.

20 Figure 2 is a flow chart of a process for locating contents in a multimedia program according to an embodiment of the present invention. Firstly, a multimedia program including a stream provided with word symbol information is obtained in step 210 (S210). The word symbol information exists in a text format, for example, in the case of a multimedia digital television program stream, the captions exist in the data stream in a text
25 format; in the case of a multimedia demonstration program stream, the wording contents for the demonstration exist in a text stream in a text format. If the multimedia program is relatively long, this step will not end until the entire locating process ends.

In this embodiment, the multimedia program about Australia Scenery is taken as an example. The program includes a text stream carrying corresponding commentary contents.

5 Then, a request containing specific word symbol, such as "Sydney Opera Theater", is received from a user (S230) ; the user expects that specific word symbol exist at certain position in the text stream and hopes to find the segment containing the specific word symbol in the multimedia program obtained in S210.

10 Next, the specific word symbol are searched in the text stream and it is judged whether they have been found appeared at a particular site in the text stream (S230). If they have not been found, then the process informs the user that the specific word symbol are not found in this multimedia program (S234) , and the entire process comes to an end. If they have been found, the process obtains the position information about the site
15 where they appeared (S238) , for example, that of the "Sydney Opera Theater" at "01: 03: 06" (hh:mm:ss) from the start of the program.

After that, the corresponding position of the specific word symbol in the video stream is determined on the basis of the particular synchronization rules of the multimedia program (S240), if the video at "01: 03: 06" (hh:
20 mm: ss) from the start of the program is found, the picture at the time often contains the scenery of the Sydney Opera Theater corresponding to the commentary. The synchronization rules for a multimedia program can be varied, and will not be elaborated here.

Finally, the video contents at the particular position are presented to
25 the user (S250), the pictures of the particular position contain the scenery of the Sydney Opera Theater that the user wants. Of course, it is possible to present to the user all the contents of the multimedia program, such as, the video/audio, image and text at this particular position; or to present

another part of them, for example, the audio only, to the user, to satisfy his individual needs.

In the presenting process of S250, it is also possible to present the video contents in the periods of time before and/or after this particular site appears. The duration of the period may be fixed a time-value by the user, or be fixed a default value by the system. The user may include a starting position information and a ending position information in the request of the S220, both of which correspond to the particular appearing site expected by the user.

Of course, in S240 of this embodiment, the position where the specific word symbol appear in the corresponding position of audio or image streams may be determined according to the synchronization rules. Since video, audio or even image is more complex than text in composition, the processes of analyzing and locating it are also much more complex than those of text. Thus it can be seen that the locating method developed in the present invention is much simpler than that of the prior method through audio/video.

In said locating process, if the specific word symbol, such as " Sydney Opera Theater " appears many times in said text stream, when S250 presents the video contents of the particular site to the user, the user is given a chance to choose whether to keep on searching. If he does, then the search will go on along the original search direction from the last found particular site until the user desired scenery is found or until the program ends. Such chance of choice can be provided in the form of a push button on the screen to prompt the user to decide whether or not to go on searching, and then the search ends in answering to the user's input information.

Figure 3 is a flow chart of a process for locating contents in a multimedia program and extracting particular segment according to another

embodiment of the present invention. Firstly, a multimedia program is obtained (S310), which includes a stream provided with word symbol information existing in an image format. For example, in respect of a multimedia demonstration program, its demonstration slides contain word symbol information contents, and exist in an image stream in an image format. If the multimedia program is relatively long, this step is one that continues until the entire locating process ends.

Table 1 is a SMIL Script of a multimedia demonstration program including a video stream and an image stream synchronized with the video stream; said image stream includes the demonstration slides and words on the slides, and these words are in an image format.

Table 1: A Multimedia Demonstration Program

It is seen from Table 1 that said image stream, having a layered structure, contains 9 sections: image1, Image2、image3、image4、image5、image6、image7、images、1mageg. Each section corresponds to one slide, that is, each section has its particular starting position and length of continuation for the reason that the video/audio generally change constantly during the demonstration process, and each slide is normally kept unchanged for a period of time.

Since it is impossible to directly conduct a textual analysis of the words existing in an image format, a certain means may be used to obtain the text information corresponding to the word symbol information in the image stream (S320). This obtainment step can be performed by the existing Optical Character Recognition (OCR) technology. Then, a request containing specific word symbol is received from the user (S330); the user expects that the specific word symbol exist at one or more position in said multimedia program stream and hopes to find and extract the segments including the specific word symbol through the request.

Next, the specific word symbol are searched in the word symbol information of the image stream and it is judged whether the specific word symbol have been found appeared at a particular site (S340). If they are not found, then the user is informed that the specific word symbol are not found in this multimedia program (S344), and the entire process ends. If they are, then the information appeared at the particular site (S350) is obtained. For example: these specific word symbol appear in the word symbol information of the image2, then the starting position and the duration of image2 are obtained.

After that, according to the synchronization rules of the particular multimedia program, determination is made of the corresponding position in the video stream of the site where the words appear (S360). At this time, the starting position and duration of the particular segment of the corresponding video stream are the same as those of image2.

Finally, the original SMIL Script is modified on the basis of the obtained starting position and duration of the particular segment to obtain a new SMIL Script (S370). This SMIL Script reflects only the segment found, thus making it possible to extract the user needed particular segment from the multimedia program. By selectively performing the modified SMIL Script, the user can directly browse the needed particular segment.

After the S360, further judgment can be made as to whether or not it is necessary to go on with the search (S380). If it is not, then the entire extracting process ends; if it is, the process will return to S340, then go ahead with the search from the last found particular site along the original search direction until a next segment or program the user wants to watch is found. The judgment can be made by automatically judging whether the multimedia program ends or not, or a decision is made by the user by prompting him to do so.

In this embodiment, in addition to said particular text information found in image2, said particular text information is also found in image5 and image8. The modified SMIL Script finally obtained is shown in Table 2. The multimedia program segments corresponding to the SMIL Script contain
5 said particular text information.

Table 2: A particular segment 1 of a multimedia program

Wherein: $T1 = t1$,

$T2 = t1 + t2 + t3 + t4$

$T3 = t1 + t2 + t3 + t4 + t5 + t6 + t7$

10 In this embodiment, the stream provided with word symbol information of the multimedia program has a layered structure. The layered structure can be presented as 9 parallel images arranged in sequence like the chapters of a book, that is, the respective layers can be mutually contained.

Since the present invention uses the streams provided with word
15 symbol information contained in the multimedia program to perform the locating and since the analysis of the word symbol information is much simpler than that of the audio/video information, the present invention frees the program producers from a lot of work and reduces the complexity of the work. It enables the users to relatively easily perform locating operation with
20 simpler and less expensive equipment. Furthermore, it also makes it possible to use voice recognition technology to convert dialogues in the audio into the text information to be used for the locating operation.

Although the present invention has been described in combination
25 with the specific embodiments, it is obvious for those skilled in the art to make various substitutions, modifications and changes on the basis of the preceding sections. Therefore, these substitutions, modifications and changes, if they do not depart from the spirit and fall within the scope of the following claims, should be included in the present invention.